ELEMENTS

In the midst of that vast and profound stirring of human minds, which we name the Renaissance, Descartes was convinced that too many people felt it beneath them to direct their efforts to apparently trifling problems. Again and again, in his Regulae ad directionem ingenii, he reverts to this theme. Intellectual mastery of mathematics, of the departments of science, of philosophy, is the fruit of a slow and steady accumulation of little insights. Great problems are solved by being broken down into little problems. The strokes of genius are but the outcome of a continuous habit of inquiry that grasps clearly and distinctly all that is involved in the simple things that anyone can understand.

I thought it well to begin by recalling this conviction of a famous mathematician and philosopher, for our first task will be to attain familiarity with what is meant by insight, and the only way to achieve this end is, it seems, to attend very closely to a series of instances all of which are rather remarkable for their banality.

1. <u>A Dramatic Instance.</u> Our first illustrative instance of insight will be the story of Archimedes rushing naked from the baths of Syracuse with the cryptic cry, "Eureka"! King Hiero, it seems, had had a votive crown fashioned by a smith of rare skill and doubtful honesty. He wished to know whether or not baser metals had been added to the gold. Archimedes was set the problem and in the bath had hit upon the solution. Weigh the crown in water! Implicit in this directive were the principles of displacement and of specific gravity.

With those principles of hydrostatics we are not directly concerned. For our objective is an insight into insight. Archimedes had his insight by thinking about the crown: we shall have ours by thinking about Archimedes. What we have to grasp is that insight 1) comes as a release to the tension of inquiry, 2) comes suddenly and unexpectedly, 3) is a function not of outer circumstances but inner conditions, 4) pivots between the concrete and the abstract, and 5) passes into the habitual texture of one's mind.

First, then, insight comes as a release to the tension of inquiry. This feature is dramatized in the story by Archimedes' peculiarly uninhibited exultation. But the point I would make does not lie in this outburst of delight but in the antecedent desire and effort that it betrays. For if the typical scientist's satisfaction in success is more sedate, his earnestness in inquiry can still exceed that of Archimedes. Deep within us all, emergent when the noise of other appetites is stilled, there is a drive to know, to understand, to see why, to discover the reason, to find the cause, to explain. Just what is wanted, has many names. In what precisely it consists, is a matter of dispute. But the fact of inquiry is beyond all doubt. It can absorb a man. It can keep him for hours, day after day, year after year, in the narrow prison of his study or his laboratory. It can send him on dangerous voyages of exploration. It can withdraw him from other interests, other pursuits, other pleasures, other achievements. It can fill his waking thoughts, hide from him the world of ordinary affairs, invade the very fabric of his dreams. It can demand endless sacrifices that are made without regret though there is only the hope, never a certain promise, of success. What better symbol could one find for this obscure, exigent, imperious drive, than a man, naked, running, excitedly crying, "I've got it".

Secondly, insight comes suddenly and unexpectedly. It did not occur when Archimedes was in the mood and posture that a sculptor would select to portray "The Thinker". It came in a flash, on a trivial occasion, in a moment of relaxation. Once more there is dramatized a universal aspect of insight. For it is reached, in the last analysis, not by learning rules, not by following precepts, not by studying any methodology. Discovery is a new beginning. It is the origin of new rules that supplement or even supplant the old. Genius is creative. It is genius precisely because it disregards established routines, because it originates the novelties that will be the routines of the future. Were there rules for discovery, then discoveries would be mere conclusions. Were there precepts for genius, then men of genius would be hacks. Indeed, what is true of discovery, also holds for the transmission of discoveries by teaching. For a teacher cannot undertake to make a pupil understand. All he can do is present the sensible elements in the issue in a suggestive order and with a proper distribution of emphasis. It is up to the pupils themselves to reach understanding, and they do so in varying measures of ease and rapidity. Some get the point before the teacher can finish his exposition. Others just manage to keep pace with him. Others see the light only when they go over the matter by themselves. Some finally never catch on at all: for a while they follow the classes but, sooner or later, they drop by the way.

Thirdly, insight is a function, not of outer circumstances, but of inner conditions. Many frequented the baths of Syracuse without coming to grasp the principles of hydrostatics. But who bathed there without feeling the water, or without finding it hot or cold or tepid? There is, then, a strange difference between insight and sensation. Unless one is deaf, one cannot avoid hearing. Unless one is blind, one has only to open one's eyes to see. The occurrence and the content of sensation stand in some immediate correlation with outer circumstance. But with insight, internal conditions are paramount. Thus, insight depends upon native endowment and so, with fair accuracy, one can say that insight is the act that occurs frequently in the intelligent and rarely in the stupid. Again, insight depends upon a habitual orientation, upon a perpetual alertness ever asking the little question, "Why?" Finally, insight depends on the accurate presentation of definite problems. Had Hiero not put his problem to Archimedes, had Archimedes not thought earnestly, perhaps desperately, upon it, the baths of Syracuse would have been no more famous than any others.

Fourthly, insight pivots between the concrete and the abstract. Archimedes' problem was concrete. He had to settle whether a particular crown was made of pure gold. Archimedes' solution was concrete. It was to weigh the crown in water. Yet if we ask what was the point to that procedure we have to have recourse to the abstract formulations of the principles of displacement and of specific gravity. Without that point, weighing the crown in water would be mere eccentricity. Once the point is grasped, King Hiero and his golden crown become minor historical details of no scientific importance. Once more the story dramatizes a universal aspect of insight. For if insights arise from concrete problems, if they reveal their value in concrete applications, none the less they possess a significance greater than their origins and a relevance wider than their original applications. Because insights arise with reference to the concrete, mathematicians invent symbols, teachers need black-boards, pupils have to perform experiments for themselves, doctors have to see their patients, trouble-shooters have to travel to the spot, people with a mechanical bent take things apart to see how they work. But because the significance and relevance of insight goes beyond any concrete problem or application, men formulate abstract sciences with their numbers and symbols, their technical terms and formulae, their definitions, postulates and deductions. Thus, by its very nature, insight is the mediator, the hinge, the pivot. It is insight into the concrete world of sense and imagination. Yet what is known by insight, what insight adds to sensible and imagined presentations, finds its adequate expression only in the abstract and recondite formulations of the sciences.

Fifthly, insight passes into the habitual texture of one's mind. Before Archimedes could solve his problem, he needed an instant of inspiration. But he needed no further inspiration when he went to offer the king his solution. Once one has understood, one has crossed a divide. What a moment ago was an insoluble problem, now becomes incredibly simple and obvious. Moreover, it tends to remain simple and obvious. However laborious the first occurrence of an insight may be, subsequent repetititions occur almost at will. This, too, is a universal characteristic of insight and, indeed, it constitutes the possibility of learning. For we can learn inasmuch as we can add insight to insight, inasmuch as the new does not extrude the old but complements and combines with it. Inversely, inasmuch as the subject to be learnt involves the acquisition of a whole series of insights, the process of learning is marked by an initial period of darkness in which one gropes about insecurely, in which one cannot see where one is going, in which one cannot grasp what all the fuss is about; and only gradually, as one begins to catch on, does the initial darkness yield to a subsequent period of increasing light, confidence, interest, absorption. Then, the infinitesimal calculus or theoretical physics or the issues of philosophy cease to be the mysterious and foggy realms they had seemed. Imperceptibly we shift from the helpless infancy of the beginner to the modest self-confidence of the advanced student. Eventually we become capable of taking over the teacher's role and complaining of the remarkable obtuseness of pupils that fail to see what, of course, is perfectly simple and obvious to those that understand.

2. Definition As every schoolboy knows, a circle is a locus of coplanar points equidistant from a center. What every schoolboy does not know is the difference between repeating that definition, as a parrot might, and uttering it intelligently. So, with a sidelong bow to Descartes' insistence on the importance of understanding very simple things, let us inquire into the genesis of the definition of the circle.

2.1 The Clue Imagine a cart-wheel with its bulky hub, its stout spokes, its solid rim.

Ask a question. Why is it round?

Limit the question. What is wanted is the immanent ground of the roundness of the wheel. Hence a correct answer will not introduce new data such as carts, carting, transportation, wheelwrights, or their tools. It will refer simply to the wheel.

Consider a suggestion. The wheel is round because its spokes are equal. Clearly, that will not do. The spokes could be equal yet sink unequally into the hub and rim. Again, the rim could be flat between successive spokes.

Still, we have a clue. Let the hub decrease to a point; let the rim and spokes thin out into lines; then, if there were an infinity of spokes and all were exactly equal, the rim would have to be perfectly round; inversely, were any of the spokes unequal, the rim could not avoid bumps or dents. Hence, we can say that the wheel necessarily is round, inasmuch as the distance from the center of the hub to the outside of the rim is always the same.

A number of observations are now in order. The foregoing brings us close enough to the definition of the circle. But our purpose is to attain insight, not into the circle, but into the act illustrated by insight into the circle.

The first observation, then, is that points and lines cannot be imagined. One can imagine an extremely small dot. But no matter how small a dot may be, still it has magnitude. To reach a point, all magnitude must vanish, and with all magnitude there vanishes the dot as well. One can imagine an extremely fine thread. But no matter how fine a thread may be, still it has breadth and depth as well as length. Remove from the image all breadth and depth, and there vanishes all length as well.

2.2 The Concepts The second observation is that points and lines are concepts.

Just as imagination is the playground of our desires and our fears, so conception is the play ground of our intelligence. Just as imagination can create objects never seen or heard or felt, so too conception can create objects that cannot even be imagined. How? By supposing. The imagined dot has magnitude as well as position, but the geometer says, Let us suppose it has only position. The imagined line has breadth as well as length, but the geometer says, Let us suppose it has only length.

Still, there is method in this madness. Our images and especially our dreams seem very random affairs, yet psychologists offer to explain them. Similarly, the suppositions underlying concepts may appear very fanciful, yet they too can be explained. Why did we require the hub to decrease to a point and the spokes and rim to mere lines? Because we had a clue - the equality of the spokes - and we were pushing it for all it was worth. As long as the hub had any magnitude, the spokes could sink into it unequally. As long as the spokes had any thickness, the wheel could be flat at their ends. So we supposed a point without magnitude, and lines without thickness to obtain a curve that would be perfectly, necessarily round.

Note, then, two properties of concepts. In the first place, they are constituted by the mere activity of supposing, thinking, considering, formulating, defining. They may or may not be more than that. But if they are more, then they are not merely concepts. And if they are no more than supposed or considered or thought about, still that is enough to constitute them as concepts. In the second place, concepts do not occur at random; they emerge in thinking, supposing, considering, defining, formulating: and that many named activity occurs, not at random, but in conjunction with an act of insight.

2.3 <u>The Image</u> The third observation is that the image is necessary for the insight.

Points and lines cannot be imagined. Neither can necessity or impossibility be imagined. Yet in approaching the definition of the circle, there occurred some apprehension of necessity and of impossibility. As we remarked, if all the radii are equal, the curve must be perfectly round; and if any radii are unequal, the curve cannot avoid bumps or dents.

Further, the necessity in question was not necessity in general but a necessity of roundness resulting from these equal radii. Similarly, the impossibility in question was not impossibility in the abstract but an impossibility of roundness resulting from these unequal radii. Eliminate the image of the center, the radii, the curve, and by the same stroke there vanishes all grasp of necessary or of impossible roundness.

But it is that grasp that constitutes the insight. It is the occurrence of that grasp that makes the difference between repeating the definition of a circle, as a parrot might, and uttering it intelligently, uttering it with the ability to make up a new definition for oneself.

It follows that the image is necessary for the insight. Inversely, it follows that the insight is the act of catching on to a connection between imagined equal radii and, on the other hand, a curve that is bound to look perfectly round.

2.4 The Question The fourth observation adverts to the question.

There is the question as expressed in words. Why is the wheel round?

Behind the words there may be conceptual acts of meaning, such as "wheel", "round", etc.

Behind these concepts there may be insights in which one grasps how to use such words as "wheel", "round", etc.

But what we are trying to get at, is something different. Where does the "Why?" come from? What does it reveal or represent? Already we had occasion to speak of the psychological tension that had its release in the joy of discovery. It is that tension, that drive, that desire to understand, that constitutes the primordial "Why?" Name it what you please, alertness of mind, intellectual curiosity, the spirit of inquiry, active intelligence, the drive to know. Under any name, it remains the same and is, I trust, very familiar to you.

This primordial drive, then, is the pure question. It is prior to any insights, any concepts, any words, for insights, concepts, words, have to do with answers; and before we look for answers, we want them; such wanting is the pure question.

On the other hand, though the pure question is prior to insights, concepts, and words, it presupposes experiences and images. Just as insight is into the concretely given or imagined, so the pure question is about the concretely given or imagined. It is the wonder which Aristotle claimed to be the beginning of all science and philosophy. But no one just wonders. We wonder about something.

2.5 Genesis A fifth observation distinguishes moments in the genesis of a definition.

When an animal has nothing to do, it goes to sleep. When a man has nothing to do, he may ask questions. The first moment is an awakening to one's intelligence. It is release from the dominance of biological drive and from the routines of everyday living. It is the effective emergence of wonder, of the desire to understand.

The second moment is the hint, the suggestion, the clue. Insight has begun. We have got hold of something. There is a chance that we are on the right track. Let's see.

The third moment is the process. Imagination has been released from other cares. It is free to cooperate with intellectual effort, and its cooperation consists in endeavouring to run parallel to intelligent suppositions while at the same time, restraining supposition within some limits of approximation to the imaginable field.

The fourth moment is achievement. By their cooperation, by successive adjustments, question and insight, image and concept, present a solid front. The answer is a patterned set of concepts. The image strains to approximate to the concepts. The concepts, by added conceptual determinations, can express their differences from the merely approximate image. The pivot between images and concepts is the insight. And setting the standard which insight, images, and concepts must meet is the question, the desire to know, that could have kept the process in motion by further queries, had its requirements not been satisfied.

2.6 Nominal and Explanatory Definition A sixth observation distinguishes different kinds of definition. As Euclid defined a straight line as a line lying evenly between its extremes, so he might have defined a circle as a perfectly round plane curve. As the former definition, so also the latter would serve to determine unequivocally the proper use of the names, straight line, circle. But, in fact, Euclid's definition of the circle does more than reveal the proper use of the name, circle. It includes the affirmation that in any circle all radii are exactly equal; and were that affirmation not included in the definition, then it would have had to be added as a postulate.

To view the same matter from another angle, Euclid did postulate that all right angles are equal. Let us name the sum of two adjacent right angles a straight angle. Then, if all right angles are equal, necessarily all straight angles will be equal. Inversely, if all straight angles are equal, all right angles must be equal. Now if straight lines are really straight, if they never bend in any direction, must not all straight angles be equal? Could not the postulate of the equality of straight angles be included in the defínition of the straight line, as the postulate of the equality of radii is included in the definition of the circle?

At any rate, there is a difference between nominal and explanatory difinitions. Nominal definitions merely tell us about the correct usage of names. Explanatory definitions also include something further that, were it not included in the definition, would have to be added as a postulate.

What constitutes the difference? It is not that explanatory definitions suppose an insight while nominal definitions do not. For a language is an enormously complicated tool with an almost endless variety of parts that admit a far greater number of significant combinations. If insight is needed to see how other tools are to be used properly and effectively, insight is similarly needed to use a language properly and effectively.

Still, this yields, I think, the answer to our question. Both nominal and explanatory definitions suppose insights. But a nominal definition supposes no more than an insight into the proper use of language. An explanatory definition, on the other hand, supposes a further insight into the objects to which language refers. The name, circle, is defined as a perfectly round plane curve, as the name, straight line, is defined as a line lying evenly between its extremes. But when one goes on to affirm that all radii in a circle are equal or that all right angles are equal, one no longer is talking merely of names. One is making assertions about the objects which names denote.

2.7 Primitive Terms A seventh observation adds a note on the old puzzle of primitive terms.

Every definition presupposes other terms. If these can be defined, their definitions will presuppose still other terms. But one cannot regress to infinity. Hence, either definition is based on undefined terms or else terms are defined in a circle so that each virtually defines itself.

Fortunately, we are under no necessity of accepting the argument's supposition. Definitions do not occur in a private vacuum of their own. They emerge in solidarity with experiences, images, questions, and insights. It is true enough that every definition involves several terms, but it is also true that no insight can be expressed by a single term, and it is not true that every insight presupposes previous insights.

Let us say, then, that for every basic insight there is a circle of terms and relations, such that the terms fix the relations, the relations fix the terms, and the insight fixes both. If one grasps the necessary and sufficient conditions for the perfect roundness of this imagined plane curve, then one grasps not only the circle but also the point, the line, the circumference, the radii, the plane, and equality. All the concepts tumble out together because all are needed to express adequately a single insight. All are coherent, for coherence basically means that all hang together from a single insight.

Again, there can be a set of basic insights. Such is the set underlying Euclidean geometry. Because the set of insights is coherent, they generate a set of coherent definitions. Because different objects of definition are composed of similar elements, such terms as point, line, surface, angle, keep recurring in distinct definitions. Thus, Euclid begins his exposition from a set of images, a set of insights, and a set of definitions; some of his definitions are merely nominal; some are explanatory; some are derived, partly from nominally and partly from explanatorily defined terms.

2.8 Implicit Definition A final observation introduces the notion of implicit definition.

D. Hilbert has worked out Foundations of Geometry that satisfy contemporary logicians. One of his important devices is known as implicit definition. Thus, the meaning of both point and straight line is fixed by the relation that two and only two points determine a straight line.

In terms of the foregoing analysis, one may say that implicit definition consists in explanatory definition without nominal definition. It consists in explanatory definition, for the relation that two points determine a straight line is a postulational element such as the equality of all radii in a circle. It omits nominal definition, for one cannot restrict Hilbert's point to the Euclidean meaning of position without magnitude. An ordered pair of numbers satisfies Hilbert's implicit definition of a point, for two such pairs determine a straight line. Similarly, a first degree equation satisfies Hilbert's implicit definition of a straight line, for such an equation is determined by two ordered pairs of numbers.

The significance of implicit definition is its complete generality. The omission of nominal definitions is the omission of a restriction to the objects which, in the first instance, one happens to be thinking about. The exclusive use of explanatory or postulational elements concentrates attention upon the set of relationships in which the whole of scientific significance is contained.

3.0 Higher Viewpoints The next significant step to be taken in working out the nature of insight is to analyze development. Single insights occur either in isolation or in related fields. In the latter case, they combine, cluster, coalesce, into the mastery of a subject; they ground sets of definitions, postulates, deductions; they admit applications to enormous ranges of instances. But the matter does not end there. Still further insights arise. The shortcomings of the previous position become recognized. New definitions and postulates are devised. A new and larger field of deductions is set up. Broader and more accurate applications become possible. Such a complex shift in the whole structure of insights, definitions, postulates, deductions, and applications may be referred to very briefly as the emergence of a higher viewpoint. Our question is, Just what happens?

Taking our clue from Descartes' insistence on understanding simple things, we select as our pilot instance the transition from arithmetic to elementary algebra. Moreover, to guard against possible misinterpretations, let us say that by arithmetic is meant a subject studied in grade school and that by elementary algebra is meant a subject studied in high school

3.1 Positive Integers A first step is to offer some definition of the positive integers, 1, 2, 3, 4, ...

Let us suppose an indefinite multitude of instances of "one". They may be anything one pleases, from sheep to instances of the act of counting or ordering.

Further, let us suppose as too familiar to be defined, the notions of "one", "plus", and "equals".

Then, there is an infinite series of definitions for the infinite series of positive integers, and it may be indicated symbolically by the following:

1	+	1	=	2
2	+	1	=	3
3	+	1	=	4
etc.,		etc.,		etc. . . . . . .

This symbolic indication may be interpreted in any of a variety of manners. It means one plus one equals two, or two is one more than one, or the second is the next after the first, or even the relations between classes of groups each with one, or two, or three, etc., members. As the acute reader will see, the one important element in the above series of definitions, is the etc., etc., etc., . . Without it, the positive integers cannot be defined; for they are an indefinitely great multitude; and it is only in so far as some such gesture as etc., etc., etc., is really significant, that an infinite series of definitions can occur. What, then, does the etc., etc., mean? It means that an insight should have occurred. If one has had the relevant insights, if one has caught on, if one sees how the defining can go on indefinitely, no more need be said. If one has not caught on, then the poor teacher has to labor in his apostolate of the obvious. For in defining the positive integers there is no alternative to insight.

Incidentally, it may not be amiss to recall what already has been remarked, namely, that a single insight is expressed in many concepts. In the present instance, a single insight grounds an infinity of concepts.

3.2 Addition Tables A second step will consist in making somewhat more precise the familiar notion of equality. Let us say that when equals are added to equals, the results are equal; that one is equal to one; and that therefore, an infinite series of addition tables can be constructed.

The table for adding 2 is constructed by adding one to each side of the equations that define the positive integers. Thus,

From the table	$2 + 1 = 3$
Adding 1	$2 + 1 + 1 = 3 + 1$
Hence, from the table	$2 + 2 = 4$

In like manner the whole table for adding 2 can be constructed. From this table, once it is constructed, there can be constructed a table for adding 3. From that table it will be possible to construct a table for adding 4, etc., etc., etc., which again means that an insight should have occurred.

Thus, from the definitions of the positive integers and the postulate about adding equals to equals, there follows an indefinitely great deductive expansion.

3.3. The Homogeneous Expansion The third step will be to venture into a homogeneous expansion. The familiar notion of addition is to be complemented by such further notions as multiplication, powers, subtraction, division, and roots. This development, however, is to be homogeneous and by that is meant that no change is to be involved in the notions already employed.

Thus, multiplication is to mean adding a number to itself so many times, so that five by three will mean the addition of three five's. Similarly, powers are to mean that a number is multiplied by itself so many times, so that five to the third will mean five multiplied by five with the result multiplied again by five. On the other hand, subtraction, division, and roots will mean the inverse operations that bring one back to the starting-point.

By a few insights, that need not be indicated, it will be seen that tables for multiplication and for powers can be constructed from the addition tables. Similarly, tables for subtraction, division, and roots can be constructed from the tables for addition, multiplication and powers.

The homogeneous expansion constitutes a vast extension of the initial deductive expansion. It consists in introducing new operations. Its characteristic is that the new operations involve no modification of the old.

3.4. The Need of a Higher Viewpoint A fourth step will be the discovery of the need of a higher viewpoint. This arises when the inverse operations are allowed full generality, when they are not restricted to bringing one back to one's starting-point. Then, subtraction reveals the possibility of negative numbers, division reveals the possibility of fractions, roots reveal the possibility of surds. Further, there arise questions about the meaning of operations. What is multiplication when one multiplies negative numbers or fractions or surds? What is subtraction when one subtracts a negative number? etc., etc., etc. Indeed, even the meaning of "one" and of "equals" becomes confused, for there are recurring decimals and it can be shown that point nine recurring is equal to one.

let	X	$=$	$0.\dot{9}$
then	$10X$	$=$	$9.\dot{9}$
hence	$9X$	$=$	9
and so	X	$=$	1

3.5 Formulation of the Higher Viewpoint A fifth step will be to formulate a higher viewpoint. Distinguish 1) rules, 2) operations, and 3) numbers. Let numbers be defined implicitly by operations, so that the result of any operation will be a number and any number can be the result of an operation.

Let operations be defined implicitly by rules, so that what is done in accord with rules is an operation.

The trick will be to obtain the rules that fix the operations which fix the numbers.

The emergence of the higher viewpoint is the performance of this trick. It consists in an insight that 1) arises upon the operations performed according to the old rules and 2) is expressed in the formulation of the new rules.

Let me explain. From the image of the cart-wheel we proceeded by insight to the definition of the circle. But, while the cart-wheel was imagined, the circle consists of a point and a line neither of which can be imagined. Between the cart-wheel and the circle there is an approximation but only an approximation. Now, the transition from arithmetic to elementary algebra is the same sort of thing. For an image of the cart-wheel one substitutes the image of what may be named "doing arithmetic"; it is a large, dynamic, virtual image that includes writing down, adding, multiplying, subtracting, and dividing numbers in accord with the precepts of the homogeneous expansion. Not all of this image will be present at once, but any part of it can be present and, when one is on the alert, any part that happens to be relevant will pop into view. In this large and virtual image, then, there is to be grasped a new set of rules governing operations. The new rules will not be exactly the same as the old rules. They will be more symmetrical. They will be more exact. They will be more general. In brief, they will differ from the old much as the highly exact and symmetrical circle differs from the cart-wheel.

What are the new rules? In high school the rules for fractions were generalized; rules for signs were introduced; rules for equations and for indices were worked out. Their effect was to redefine the notions of addition, multiplication, powers, subtraction, division, and roots; and the effect of the redefinitions of the operations was that numbers were generated, not merely by addition, but by any of the operations.

3.6 Successive Higher Viewpoints The reader familiar with group theory will be aware that the definition of operations by rules and of numbers or, more generally, symbols by operations is a procedure that penetrates deeply into the nature of mathematics. But there is a further aspect to the matter, and it has to do with the gradual development by which one advances through intermediate stages from elementary to higher mathematics.

The logical analyst can leap from the positive integers to group theory, but one cannot learn mathematics in that simple fashion. On the contrary, one has to perform, over and over, the same type of transition as occurs in advancing from arithmetic to elementary algebra.

At each stage of the process there exists a set of rules that govern operations which result in numbers. To each stage there corresponds a symbolic image of doing arithmetic, doing algebra, doing calculus. In each successive image there is the potentiality of grasping by insight a higher set of rules that will govern the operations and by them elicit the numbers or symbols of the next stage. Only in so far as a man makes his slow progress up that escalator does he become a technically competent mathematician. Without it, he may acquire a rough idea of what mathematics is about; but he will never be a master, perfectly aware of the precise meaning and the exact implications of every symbol and operation.

3.7 The Significance of Symbolism The analysis also reveals the importance of an apt symbolism.

There is no doubt that, though symbols are signs chosen by convention, still some choices are highly fruitful while others are not. It is easy enough to take the square root of 1764. It is another matter to take the square root of MDCCLXIV. The development of the calculus is easily designated in using Leibniz' symbol dy/dx, for the differential coefficient; Newton's symbol, on the other hand, can be used only in a few cases and, what is worse, it does not suggest the theorems that can be established.

Why is this so? It is because mathematical operations are not merely the logical expansion of conceptual premises. Image and question, insight and concepts, all combine. The function of the symbolism is to supply the relevant image, and the symbolism is apt inasmuch as its immanent patterns as well as the dynamic patterns of its manipulation run parallel to the rules and operations that have been grasped by insight and formulated in concepts.

The benefits of this parallelism are manifold. In the first place, the symbolism itself takes over a notable part of the solution of problems, for the symbols, complemented by habits that have become automatic, dictate what has to be done. Thus, a mathematician will work at a problem up to a point and then announce that the rest is mere routine. In the second place, the symbolism constitutes a heuristic technique; the mathematician is not content to seek his unknowns; he names them; he assigns them symbols; he writes down in equations all their properties; he knowns how many equations he will need; and when he has reached that number, he can say that now the solution is automatic.
In the third place, the symbolism offers clues, hints, suggestions. Just as the definition of the circle was approached from the clue of the equality of the spokes, so generally insights do not come to us in their full stature; we begin from little hints, from suspicions, from possibilities; we try them out; if they lead nowhere, we drop them; if they promise success, we push them for all they are worth. But this can be done only if we chance upon the hints, the

clues, the possibilities; and the effect of the apt symbolism is to reduce, if not entirely eliminate, this element of chance. Here, of course, the classical example is analytic geometry. To solve a problem by Euclidean methods, one has to stumble upon the correct construction. To solve a problem analytically, one has only to manipulate the symbols.

In the fourth place, there is the highly significant notion of invariance. An apt symbolism will endow the pattern of a mathematical expression with the totality of its meaning. Whether or not one uses the Latin, Greek, or Hebrew alphabet, is a matter of no importance. The mathematical meaning of an expression resides in the distinction between constants and variables and in the signs or collocations that dictate operations of combining, multiplying, summing, differentiating, integrating, and so forth. It follows that, as long as the symbolic pattern of a mathematical expression is unchanged, its mathematical meaning is unchanged. Further, it follows that if a symbolic pattern is unchanged by any substitutions of a determinate group, then the mathematical meaning of the pattern is independent of the meaning of the substitutions.

In the fifth place, as has already been mentioned, the symbolism appropriate to any stage of mathematical development, provides the image in which may be grasped by insight the rules for the next stage.

4. Inverse Insight So far we have been asking questions that can be answered. How can one tell whether a crown is made

of pure gold without melting it down? What accounts for a wheel being round? What is arithmetic and how does one go on to algebra? In each case, there is an appropriate image or set of images that, under the stress of inquiry, results in an insight that expresses itself in some formulation called the answer.

Now attention has to be directed to a quite different case. There is the question. There is the answer. But the answer consists in showing the question to be misconceived, and it is grounded in an insight that grasps why the question, as conceived, cannot be answered.

4.1 Surds How big is the square root of two? Clearly, it is greater than one, for the square of one is one; and it is less than two, for the square of two is four. It would seem, then, that it is some improper fraction lying between one and two.

Now an improper fraction is the quotient of some positive integer divided by some other, smaller positive integer. Moreover, it is always possible to reduce such a fraction to its lowest terms by removing all common factors. Let us suppose then, that:

$$\sqrt{2} = M/N$$

where M and N are positive integers with no common factors. Multiplying across by N and squaring, one obtains:-

$$2N^2 = M^2$$

It follows that M must be an even number and so twice, say, P.

Substituting and dividing by two, one obtains:

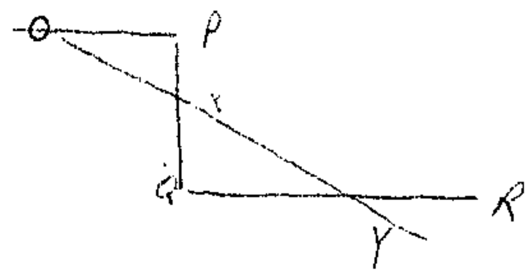
$$N^2 = 2P^2$$

so that N also must be an even number, which contradicts the assumption that all common factors were eliminated. It follows that there is no "rational" fraction, M/N, that is equal to the square root of two. Moreover, since any recurring decimal can be reduced to such a fraction, there is no recurring decimal equal to the square root of two. However, one can apply to 2 the ordinary method for taking the square root, and so it remains that the square root of two will be an infinite, non-recurring decimal. Finally, the foregoing argument can be generalized and applied to any surd. Thus, if

$$3N^2 = M^2$$

then 3 must be a factor of M, so that M can be replaced by 3P, whence, it will follow that 3 must be a factor of N.

4.2 Non-Countable Infinity Again, to raise another, similar question. How many points are there in a straight line one inch long? Clearly, the number must be very large, for a point is position without magnitude. But, at least, one would be inclined to say that there are twice as many points in a straight line two inches long as in a straight line one inch long. Still, that would be erroneous, as appears from the following construction. Let the straight line, PQ, be perpendicular to the straight lines, OP and QR. Let QR be twice as long as PQ and let OXY be a straight line cutting PQ in X and QR in Y.



Then, from the construction, it is clear that for every point, Y, in QR, there is a corresponding, distinct point, X, in PQ. Indeed, this remains true if one produces QR to infinity in the direction of R. No matter where Y is taken on QR produced, there is always a corresponding and distinct point X in PQ. Hence, there are as many points in an inch of straight line as there are in two inches, or in a foot, or in a mile, or in as many light-years as you please.

However, we have not met the question. We have said there are <u>as many as</u>. . . . We have not said <u>how many</u>. Accordingly, let us distinguish between the counted, the countable, and the non-countable. A set is counted when one says it contains N members, where N is some positive integer. A set is countable when it can be arranged in some determinate order that contains all its members once each and only once; for then there can be established a one-to-one correspondence between the members of the set and the positive integers. Finally, a set is non-countable when it is not possible to establish a one-to-one correspondence between its members and the positive integers.

It is to be noted that by "countable" is not meant the possibility of finishing the counting. Thus, an infinite series, such as

1/2, 1/4, 1/8, 1/16, ...

is countable, for all its members lie in a determinate order and so can be placed in a one-to-one correspondence with the positive integers. Again, an infinite series of infinite series of elements is countable, for all its elements can be regarded as lying within a single determinate order. Thus, the reciprocals of the <u>n</u>th powers of the prime numbers form an infinite series of infinite series. Their elements can be arranged in a column of rows, thus:

1/2	1/4	1/8	1/16	. . . . .
1/3	1/9	1/27	1/81	. . . . .
1/5	1/25	1/125	1/525	. . . . .
1/7	1/49	1/343	1/2401	. . . . .

and any column of rows can be counted in the following manner:

1	2	5	10	17
4	3	6	11	etc. .
9	8	7	12	etc. .
16	15	14	13	etc. .

Thus, any infinite series of infinite series can be assigned the order of a single infinite series. It follows that an infinite series of infinite series of infinite series can be arranged in a column of rows and so can be assigned the order of a single infinite series. The theorem can be repeated indefinitely. Thus, consider the rational, proper fractions:

1/2, 1/3, 2/3, 1/4, 3/4, 1/5, 2/5, 3/5, 4/5, 1/6, ...

From this infinite series there can be derived an infinite series of infinite series, for one can take, first, the

Found thus, Apr. 5/'29. Of no signif. F.C.

square root of the lot, then, the cube root, then square the cube roots, then take the fourth root, then cube the fourth roots, etc.... Now, as has been shown, this infinite series of infinite series can be arranged in a single series. Once this is done, one can use these new terms as powers to be applied to the rational proper fractions to derive a new infinite series of infinite series. This can be arranged in a single series, applied as powers to the rational proper fractions, yield a new infinite series of infinite series, etc., etc.

From the foregoing it is clear that any infinite set is countable, provided it is possible to assign some order to its members. It is also clear that a non-countable infinite set must contain such a multitude of members in such a manner that ordering them is impossible. Such is the case with the points in a straight line. Thus, in the line, QR, it is impossible to pick any point, Q', that is nearest to Q; for however short QQ' may be, it contains as many points as there are in a line as long as you please. Nor is there any use trying to proceed by dividing the line. For if this could be done in an orderly fashion, then one would be appealing to an ordered series of all the numbers greater than zero and less than unity. But the range of numbers is a non-countable infinite set, for it cannot be arranged in a single order. Suppose there were some single column containing all the infinite decimals. Then consider the diagonal. It is always possible to construct another infinite decimal that differs from the first infinite decimal by the digit in the first place, from the second by the digit in the second place, from the nth by the digit in the nth place. Therefore, the initial assumption is false. The column did not contain all the infinite decimals. There is, then, no single series that contains all the infinite decimals and so the infinite decimals are a non-countable infinite set.

Well, how many points are there in a straight line an inch long? There is no answer. They form a non-countable infinite set. They do so, because they cannot be placed in a single order and so cannot be correlated in a one-to-one correspondence with the positive integers. However, they can be placed in a one-to-one correspondence with other non-countable infinite sets. Thus, there are as many points in an inch as in a mile or in a light-year or in as many light-years as you please. But that does not mean that there is some determinate number of points in an inch or in a mile. Much less does it mean that some smaller number is equal to a greater number. There just is no numbering, no counting. And there is no numbering or counting because there is no possibility of effecting an order, a system, an arrangement.

4.3 Function and Limit One might think that this exclusion of number and of order blocked the mathematician. In fact, it gives him a new lease of life. What is the mathematician's continuous function? In the elementary case, it is a one-to-one correspondence between non-countable infinite sets. Moreover, since such a correspondence can be set up between an inch and a foot, or an inch and a mile, or an inch and a light-year, or any intermediate or still odder pair, since, visually, length is independent of the number of points, the mathematician can develop the infinitesimal calculus. But he does so, not by finding some order in the non-countable infinite set, but by developing a technique of getting around it. This technique is named proceeding to the limit.

Thus, consider the continuous function, $y = x^2$. It is a function if, for every value of x, there is a corresponding value of y. It is a continuous function, if the values of x are a non-countable infinite set.

Now as x increases, y must increase more rapidly, for it equals the square of x. Hence, visually, as one moves from point to point along x, one must move more rapidly from point to point along y. Moreover, the further one advances along x, the greater must be one's strides along y. Still, there are no points omitted along x and there are no points omitted along y.

What, then, is the ratio of the increment of y to the increment of x? Clearly, if x increases by some slight amount, h, y will increase by

$$(x+h)^2 - x^2 = 2xh + h^2$$

Hence the ratio of the corresponding increment of y to the increment, h, of x will be $(2x+h)$. The smaller the increment, h, the nearer is the ratio to 2x. In the limit, it is exactly 2x. Thus, if the limit of the ratio of the increment in y to the increment in x is denoted by the symbol, dy/dx, then besides the initial function, $y = x^2$, we have also the derivative function, $dy/dx = 2x$.

Now what is this business of proceeding to the limit? There is said to be a limit, P., to a non-determined quantity, Q, if the difference of Q from P can be made smaller than any number one cares to assign. Thus, above, by making the increment, h, smaller and smaller, one can make the difference of $(2x + h)$ from $2x$ as small as we please. Still, this is only the conceptual formulation of the procedure of taking a limit. What is the underlying insight? What is the image that the insight presupposes?

Clearly enough the image will differ in different cases. Similarly, the insight will be reached in different manners. But the peculiarity of the insight is that it grasps, not that something is to some point, but that something is beside the point. No matter how small h is, there is a non-countable infinite set of values between $2x$ and $(2x + h)$. They are non-countable because they defy arrangement, order, system. They exhibit a non-systematic aspect of continuous variables and continuous functions. But what one is trying to do in mathematics, is to reach the systematic. If that is all one wants, one can disregard the non-systematic. One can leap over the non-countable infinity because it is without order if one's aim is to grasp just what admits order. Again, the ratio of the increment of y to the increment of x is any of a non-countable infinite set of values. But the limit of that ratio is unique. It can be determined systematically. It pertains to system.

4.4 Abstraction When one comes to think of it, we have been doing this sort of thing all along. The principles of displacement and of specific gravity would not enable Archimedes to determine that there was nothing but pure gold in the crown; they would enable him to say merely that there was extremely little else. Again, the definition of the circle paid no attention to the size, the weight, the strength, the origin, the materials, the purpose of the cart-wheel; on the contrary, it went off to a realm of the non-imaginable where points have position without magnitude and lines have length without thickness. Finally, the transition from arithmetic to algebra did not consist in paying closer attention to the things one might count by the positive integers; it consisted in deserting the good, common sense notion of adding and in developing a new notion that gave a meaning to adding negative numbers, multiplying fractions, and doing other things that have no prima facie meaning.

It is time, then, for us to reflect on certain general aspects of the process from image through insight to conceptions, and so we had best begin a new section.

5. The Empirical Residue The suppositions and conceptions resulting from insight commonly are abstract. They abstract from the irrelevant, the insignificant, the negligible, the incidental. They concentrate upon the relevant, the significant, the important, the essential.

But what is the relevant, the significant, the important, the essential? The answer depends immediately upon the insight, or set of insights, grounding the supposing, considering, thinking, defining, formulating. Ultimately, one will say that the answer depends on which insight, or set of insights, is right. But we are not yet ready to tackle ultimate questions. Accordingly, we have to acknowledge, for the present, that the relevant and the irrelevant, the significant and the insignificant, the important and the negligible, the essential and incidental, vary with one's insights. What at one time, one thinks important, later, in the light of fuller insight, one will think unimportant. Inversely, what one used to think insignificant, now one may think significant; and what the difference is the advent of further insight.

Still, even for the present, this relative pronouncement is not the whole story. For if we restrict ourselves to the insights possible in mathematics, physics, chemistry, biology, sensitive psychology, and such sciences, then there are elements or components in sensible data and in images that always are regarded as irrelevant, insignificant, negligible, incidental. Such elements or components may be named the empirical residue. They are given as a matter of fact. But they are always disregarded when one concentrates on whatever one happens to think essential.

On four aspects of this empirical residue, something must now be said. They are 1) the individual, 2) the continuum, 3) place and time, and 4) the actual frequency of events.

5.1 Individuality Individuality pertains to the empirical residue. For whenever we understand anything, we would understand an exactly similar instance in exactly the same fashion. A different understanding would presuppose a difference in the data. It would presuppose the possibility of saying that the previous understanding would do, were it not for this aspect of the object. But, *ex Hypothesi*, there is no aspect in which the second object differs from the first, and so there is no possibility of a different understanding. One may learn something new when one turns to the second object; but one automatically learns it about the first object as well.

Thus, a first motor car off the assembly line may be understood in terms of certain principles of construction and of operation. A second motor car, similar in all respects, cannot but be understood in exactly the same fashion.

Nor is the issue changed essentially when one understands instances that are unique. In this case, there is no possibility of apprehending a second object and understanding it in the same manner. But there is the possibility of apprehending the same object a second time; the data in the second apprehension will be similar to those of the first; because the data are similar, the understanding has to be the same. The fact that the similar data are of the same object does not alter the underlying principle that our knowledge is so constituted that similar data have to result in similar insights with the consequence that, what is grasped by insight, is independent of the individuality of the data.

Thus, if the development of all life on this planet were comprehended in a single evolution, there would be no remainder of life on the planet to be understood in either the same or a different fashion. The object would be unique and unparalleled in our experience. None the less, the understanding would consist in grasping principles and laws in the combinations suitable for mastering the enormous ranges of data, while knowledge of the unique instance would consist in observing the data to be understood.

Again, what is grasped by insight, may be named an idea or form emergent in sensible presentations or imaginative representations. But it is one thing to say that grasp of such an idea or form is knowledge of individuality, and quite another to say that within our experience there is found only one instance in which the idea or form can be grasped. If grasp of the idea or form were knowledge of individuality, then the individual would be known by understanding and it would not pertain to the empirical residue. But the mere fact that in some cases there is but a single, observable instance, in which the idea or form can be grasped, provides no evidence for the intrinsic intelligibility of individuality.

In brief, nothing is explained by saying that it is this instance. Inversely, in so far as we grasp explanations, we know not instances but what may or may not be found in individual instances.

5.2 Continuity The continuum pertains to the empirical residue. Let us begin with a definition that departs from ordinary mathematical usage to meet our present purpose. A variable, x will be said to be continuous in the range, $a < x < b$, if the values of x in every part of the range form non-countable infinite sets. Next, a function, f(x), will be said to be continuous in a range if 1) x is continuous in the range and 2) for every distinct value of x there is a corresponding value of the function. Finally, continuous functions possess a number of distinctive properties; hence, through the verification of these distinctive properties, it may be possible to verify the existence of continuous functions and so conclude to the existence of continua.

Now a continuum, in this defined and verifiable sense (which does not suppose a non-countable infinite set of observations) includes what cannot be counted because it cannot be ordered or systematized. By this inclusion of the non-systematic, a continuum clearly pertains to the empirical residue.

5.3 Place and Time Place and time pertain to the empirical residue. Space is a continuum of individual positions. Time is a continuum of individual instants. No position is any other. No instant is any other. And of both there are non-countable infinite sets. But the individual and the continuum both pertain to the empirical residue. So also, then, must place and time in their basic aspects.

Hence, when different experimenters, performing the same experiment at different places or times, obtain different results, then no one dreams of explaining the difference in the results by the differences in the place or by the difference in the time. The appeal always is, not to the place, but to something in the place, and not to the time, but to something at the time.

Indeed, if place or time made any difference, then each place and each time would have its own physics, chemistry, and biology. For if place were relevant, the laws in one place could not be the laws in another. If time was relevant, the laws at one time could not be the laws at another. Further, since places and times are non-countable sets, there would be non-countable sets of different physics, different chemistries, different biologies. Finally, none of the elements of these sets could be ascertained. For one cannot set up a whole physics, or a whole chemistry, or a whole biology, with the observations or experiments made at a point-instant.

However, it is only in their basic aspects that place and time pertain to the empirical residue. A place can be of singular importance, provided that importance rests not on a mere "there" but on a "something there". Such is the importance of the place occupied by a central mass in a gravitational field. Similarly, a time can be of singular importance, provided its importance rests not on a mere "then" but on what happened then. Such is the importance of the initial moment in certain theories of the expanding universe.

5.4 Actual Frequency Actual frequency pertains to the empirical residue.

The probability of tossing "heads" is 1/2. But in any series of actual tosses, one does not obtain a regular alternation of "heads" and "tails". Between the probability and the actual frequency, there is a divergence. Moreover, this divergence is random. It cannot be reduced to any law or mitigated by any reasonable expectation. It is non-systematic. It is to be known in each case only by actual observation. It too pertains to an empirical residue.

5.5 The Significance of the Empirical Residue

Let us now recall an initial restriction. The empirical residue was defined as always irrelevant from the viewpoint of insights in mathematics and natural sciences. Why was this restriction imposed? Quite clearly, because in such a science as the theory of knowledge the notion of the empirical residue attains a systematic significance. For in a study of knowledge one attends systematically, not only to what is concentrated upon in abstraction, but also to what is regularly abstracted from. Theory of knowledge is a higher level science that takes as its materials the whole of the knowledge in other sciences.

Indeed, the theoretical account of the empirical residue is of considerable significance.

It is because insight abstracts from the individual that science is of the universal. It is because science is of the universal, that the observation of similarities is of such great heuristic importance.

It is because insight abstracts from the continuum by proceeding to the limit that the infinitesimal calculus is such a unique and powerful instrument in the construction of theories.

It is because insight abstracts from place and time that principles and laws are independent of place and time and that the expression of principles and laws is invariant with respect to transformations of certain groups of coordinate systems.

It is because insight abstracts from the random divergence of the actual frequency that probability theory has its place among the instruments of scientific knowledge.

Generally, corresponding to each aspect of the empirical residue, there will be a remarkably powerful technique of intelligence in mastering the multiplicity of sensible data. Unfortunately, the discovery of the techniques has to be prior to the determination of the complementary aspect of the empirical residue. For while all aspects of the empirical residue are given on the level of observation, still one can grasp them as pertaining to the empirical residue only by understanding the corresponding techniques.